

Systematic Review

Psychometric properties of measures designed to assess common mental health problems and wellbeing in adults with intellectual disabilities: a systematic review

M. Patel,  J. Y. Lee  & K. Scior *Research Department of Clinical, Educational and Health Psychology, University College London, London, UK*

Abstract

Background Multiple measures of mental health problems and mental wellbeing for adults with intellectual disabilities are available, but investigations into their reliability and validity are still in the early stages. The aim of this systematic review was to provide an update to previous evaluations of measures of common mental health problems and wellbeing in adults with mild to moderate intellectual disabilities (ID).

Methods A systematic search was performed across three databases (MEDLINE, PsycINFO and SCOPUS). The literature search was limited to the years from 2009 to 2021 and to the original English versions. Ten papers evaluating nine measures were reviewed, and the psychometric properties of these measures were discussed using the Characteristics of Assessment Instructions for Psychiatric Disorders in Persons with Intellectual Developmental Disorders as a framework.

Results Four measures had at least one rating of 'good' across both dimensions of reliability and at least one dimension of validity and were deemed to

have promising psychometric properties: the Clinical Outcomes in Routine Evaluation–Learning Disabilities, Impact of Events Scale–Intellectual Disabilities, Lancaster and Northgate Trauma Scales and Self-Assessment and Intervention (self-report section). Additionally, these measures were developed through consultations with mental health professionals and/or people with IDs and thus were deemed to have good content validity.

Conclusions This review informs measurement choice for researchers and clinicians while highlighting a need for continued research efforts into the quality of measures available for people with IDs. The results were limited by incomplete psychometric evaluations of measures available. A paucity of psychometrically robust measures of mental wellbeing was observed.

Keywords intellectual disability, measurement, mental health, mental wellbeing, psychometric properties

Introduction

People with IDs may experience higher rates of mental health problems compared with the general population (Cooper *et al.* 2007; Dunn *et al.* 2020),

Correspondence: Dr Maya Patel, Research Department of Clinical, Educational and Health Psychology, University College London, Gower London WC1E 6BT, UK (e-mail: maya.patel.12@ucl.ac.uk).

© 2023 The Authors. Journal of Intellectual Disability Research published by MENCAP and International Association of the Scientific Study of Intellectual and Developmental Disabilities and John Wiley & Sons Ltd.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

although estimates of prevalence differ by sampling method and population studied ranging from 10% to 39% (Emerson & Hatton 2007; Pouls *et al.* 2021). The most common mental health problems, depression and anxiety, have been estimated to affect between 2.2% and 15.8%, and 3.8% and 17.4%, respectively, of this population (Reid *et al.* 2011; Hsieh *et al.* 2020).

Although there is a range of both global- and symptom-specific measures of mental health problems for adults with intellectual disabilities (IDs) available, investigations into their reliability and validity are still in the early stages (Hatton & Taylor 2013). To date, two systematic reviews of measures of depressive symptoms in people with IDs have been conducted (Perez-Achiaga *et al.* 2009; Hermans & Evenhuis 2010). The earlier review concluded that the Reiss Screen for Maladaptive Behaviour (RSMB; Reiss 1988) and the Psychiatric Assessment Schedule for Adults with Developmental Disabilities Checklist (PAS-ADD; Moss *et al.* 1993), two global mental health measures, demonstrated robust psychometric properties. However, Hermans & Evenhuis (2010) disagreed on the utility of the PAS-ADD Checklist and the RSMB for screening for depression, as sensitivity and specificity had not been measured, although agreed that the psychometric properties of the RSMB were promising. They concluded that the Glasgow Depression Scale for people with a Learning Disability (GDS-LD; Cuthill *et al.* 2003) was the most promising self-report instrument, while the Assessment of Dual Diagnosis (ADD; Matson & Bamberg 1998), the RSMB and the Children's Depression Inventory (CDI; Kovacs 1985) were promising informant-report measures. However, they noted that none of these informant-report measures had yet been satisfactorily assessed with regard to their psychometric properties when used with this population. Furthermore, Hermans *et al.* (2011) conducted a systematic review of measures of anxiety for people with IDs. They concluded that the Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID; Mindham & Espie 2003) was the most robust self-report instrument, whereas the Anxiety, Depression And Mood Scale (ADAMS; Esbensen *et al.* 2003) was the most promising informant-report instrument.

In the field of mental health, there is a growing interest in promoting 'positive mental health', or

mental wellbeing. Several conceptualisations of mental wellbeing have been debated, although the consensus is that wellbeing encompasses 'feeling well' (hedonia) and 'functioning well' (eudaimonia), as opposed to a mere absence of symptoms of mental illness (Keyes 2002; Deci & Ryan 2008; Stewart-Brown *et al.* 2015; Cooke *et al.* 2016). Evidence suggests that mental wellbeing may protect against psychopathology (Trompeter *et al.* 2017). Compared with the general population, there is less research pertaining to individuals with IDs in this area (Raczka *et al.* 2020).

The term 'mental wellbeing' has often been used interchangeably with 'quality of life' (QoL) in the literature (Cooke *et al.* 2016) although it has been argued that they refer to different theoretical concepts (Skevington & Böhnke 2018), with QoL referring to 'an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns' (WHOQOL Group 1995, p. 1404). Two systematic reviews to date (Townsend-White *et al.* 2012; Li *et al.* 2013) have explored the measurement of QoL in adults with IDs, and these reviews included search terms related to 'wellbeing'. In the first review, the authors concluded that the Choice Questionnaire (CQ; Stancliffe & Parmenter 1999) and the Quality of Life Questionnaire (QOLQ; Schalock & Keith 1993) were the most psychometrically robust measures, while the authors of the subsequent review concluded that 6 out of the 24 measures of QoL they evaluated were psychometrically sound, although they did not express a preference for a particular measure. Flynn *et al.* (2017) recently conducted a systematic review of measures of mental health problems and mental wellbeing in children and adults with severe or profound IDs. The Aberrant Behaviour Checklist (ABC; Aman *et al.* 1985), the Diagnostic Assessment for the Severely Handicapped Scale-II (DASH-II; Matson 1995) and the Mood, Interest and Pleasure Questionnaire (MIPQ; Ross & Oliver 2002, 2003) were rated as having good methodological quality for use with individuals who had severe to profound IDs. The authors noted that tools measuring mental wellbeing in this population were lacking.

The aim of this paper is to extend the results of the aforementioned systematic reviews and provide an update to previous psychometric evaluations of

measures of mental health problems and wellbeing, in adults with IDs. This will inform choice for clinicians and researchers interested in assessing mental health problems and mental wellbeing in this population.

As Flynn *et al.* (2017) recently evaluated measures of mental health problems and mental wellbeing for people with severe or profound IDs, this paper will evaluate measures used for people with mild to moderate IDs. In addition to mental wellbeing, the present review will focus on the measurement of anxiety disorders or depression, described by NICE (2011) as ‘common mental health problems’, because when combined, they affect more people than other mental health problems.

The review sets out to answer the following questions:

- 1 Which measures have been used to assess common mental health problems and mental wellbeing in adults with mild to moderate intellectual disabilities?
- 2 What are the psychometric properties of these measurement tools?

Method

Design

The protocol for the present review was registered with Prospero (<https://www.crd.york.ac.uk/prospero/>; registration number: CRD42021270069).

The PICO (Population, Intervention, Comparator, Outcome) framework was considered to guide the search strategy. The population was adults (aged 18 + years) with a mild to moderate ID. The intervention was the psychometric evaluation of measures. With regard to the comparator, an evaluative and descriptive tool, the Characteristics of Assessment Instructions for Psychiatric Disorders in Persons with Intellectual Developmental Disorders (CAPs-IDD; Zeilinger *et al.* 2013), was used to allow comparison between the measures. The outcomes of interest were symptoms of anxiety disorders and depression, and mental wellbeing.

Search strategy

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines

were used to inform the methodology employed. Electronic searches of the databases MEDLINE, PsycINFO and SCOPUS were conducted on 13 September 2021. The searches were limited to papers published from January 2009 to September 2021, in order to minimise overlap with previous reviews (Hermans & Evenhuis 2010; Hermans *et al.* 2011; Li *et al.* 2013), while providing an update to the literature. The full list of search terms used is summarised in Table 1. Search terms were identified based on previous similar reviews (Hermans & Evenhuis 2010; Flynn *et al.* 2017) and related to four headings, truncated where appropriate and combined using the Boolean terms ‘OR’ and ‘AND’. These were as follows: (1) psychometric properties (e.g. validity, reliability and quality); (2) measurement (e.g. assessment, outcome, screening and questionnaire); (3) common mental health problems and mental wellbeing (e.g. anxiety, depression, mood and quality of life); (4) IDs (e.g. learning disability and intellectual developmental disorder). As the focus of the present review was measures used with adults, an additional term, ‘NOT’, was used to exclude papers related to children and adolescents. Additional synonyms of these headings were also used and search terms accounted for both British English and American English spelling.

Inclusion and exclusion criteria

Papers were included if they meet the following:

- 1 The participants in the study were adults aged 18 + years old. If a study included any participants who were aged 17 years or below, the paper was included if the results for the participants aged above and below 18 years were reported separately.
- 2 At least 50% of the sample were reported to have a mild to moderate ID, defined using administrative definitions (e.g. use of specialist services for people with ID). This was to ensure that there was a majority of people with mild to moderate IDs in the study sample.
- 3 The measure assessed common mental health problems or mental wellbeing. NICE (2011) identified depression and anxiety disorders including generalised anxiety disorder, panic disorder, specific phobias, obsessive–compulsive disorder,

Table 1 List of terms used in systematic search

Psychometric properties	Measures	Common mental health problems and wellbeing	Intellectual disabilities	NOT
valid* or properties or quality or methodology* or reliab* or feasib* or psychometr* or sensitiv* or specificity	assess* or outcome* or index* or inventor* or item* or measur* or subscale* or scale* or screen* or tool* or survey* or self report* or test* or rating* or score* or questionnaire*	Mental health or mood or depress* or anxiet* or anxioi* or phobia or panic or traum* or gad or worry or worrie* or sadness or post traumatic or ptsd or ocd or obsessive compuls* or body dysmorph* psychosocial or wellbeing or well being or quality of life or happiness or life satisfaction or personal growth or self accept* or self actuali* exp Mental Health/ or exp Happiness/ or exp Well Being/or exp Emotional States/	exp intellectual development disorder/ (intellectual or development* or learning*) adj2 (disabilit* or impair* or difficult* or disorder* or handicap*) mental* retard* down* syndrome	Infant* or infancy or Newborn* or Baby* or Babies or Neonat* or Preterm* or Prematur* or Postmatu* or Child* or Schoolchild* or School age* or Preschool* or Kid or kids or Toddler* or Adolescen* or Teen* or Boy* or Girl* or Minors or Puberty or Pubescen* or Prepubescen* or Paediatric* or Pediatric* or Nursery or Kindergarten* or Primary school* or Secondary school* or Elementary school* or High school* or Highschool*

body dysmorphic disorder, post-traumatic stress disorder (PTSD), health anxiety or social anxiety as common mental health problems. Measures with items, which mapped onto several mental health problems were included if they contained items which specifically measured symptoms of anxiety disorders and/or depression, without necessarily being specific or exclusive to symptoms of anxiety and depression. Mental wellbeing was conceptualised as encompassing dimensions of hedonia (life satisfaction and positive affect) and eudaimonia (personal growth and self-acceptance). As previously mentioned, although it has been argued that they are theoretically different, the terms mental wellbeing and 'quality of life' have often been used interchangeably in the literature (Cooke *et al.* 2016). Therefore, to ensure that all of the relevant papers were included, studies, which referred to measures of either mental wellbeing or QoL, were included.

- 4 The main aim of the study was to evaluate the psychometric properties of a measure.
- 5 The article was published in a peer-reviewed journal in the English language.
- 6 The measure was administered in the English language. This was so that the review may inform measurement choice for fellow English-speaking researchers and clinicians.

Papers were excluded if the study sample included individuals aged 17 years or below and the result for these participants were not reported separately. Additionally, measures of health-related QoL were excluded due to the narrow focus on physical health-based constructs, which does not capture broader aspects of QoL or mental wellbeing.

Screening process

See Figure 1 for the PRISMA flow diagram, which summarises the systematic review screening process. The initial search yielded 3936 papers, which reduced to 2434 following the removal of duplicates. The titles were initially screened by the first author (M. P.), and if these were vague, abstracts were also reviewed. Additionally, citation searches and an inspection of reference lists were undertaken to ensure that no further eligible studies were missed. The full text of

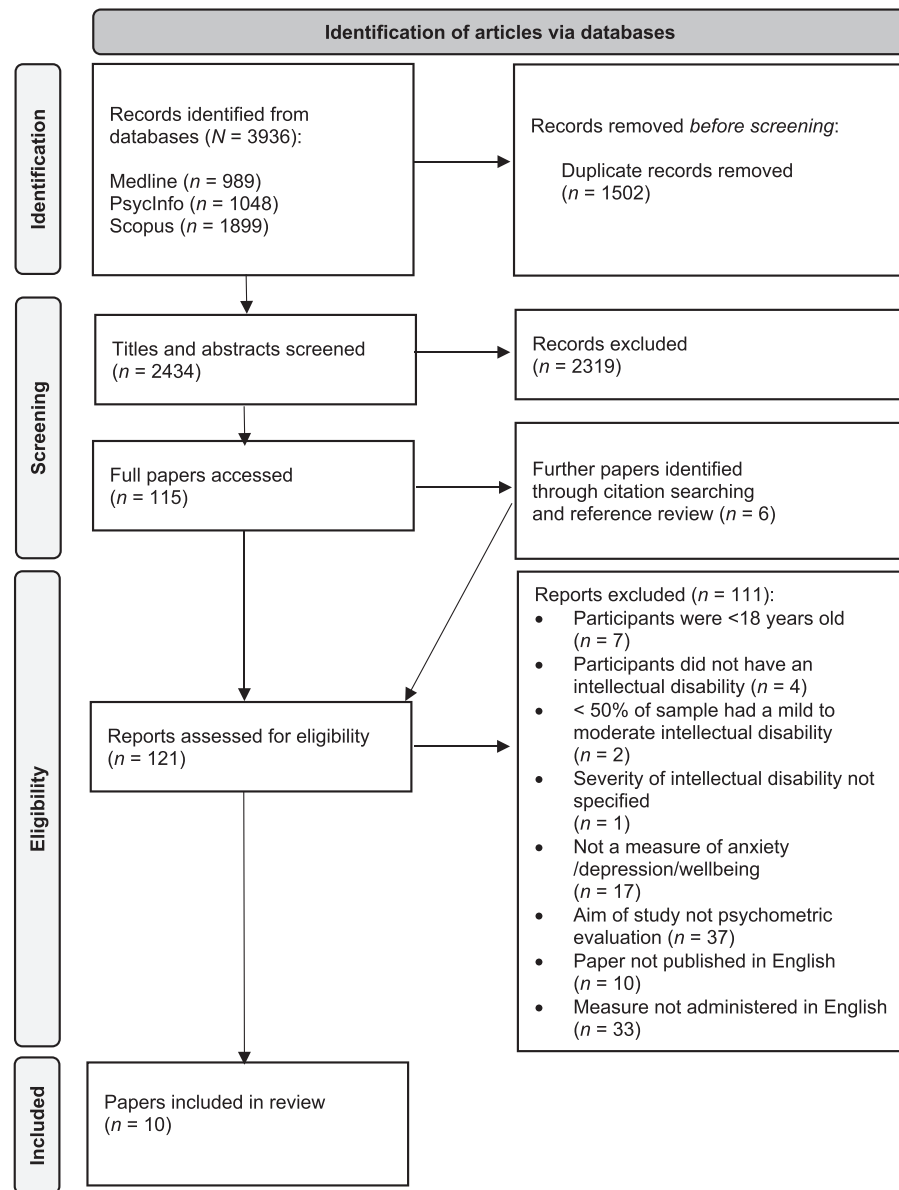


Figure 1. PRISMA flow diagram of study selection for review

potentially eligible articles ($n = 121$) was reviewed against the inclusion criteria. The third author (K. S.) independently reviewed all 121 articles against the criteria and agreed with the inclusion/exclusion of 120 papers, with one paper requiring further careful discussion with M. P. before its inclusion. Ten eligible articles, which evaluated nine measures, were included in the review.

Quality appraisal

The CAPs-IDD is a comprehensive framework for evaluating and describing measures of psychiatric disorders in people with IDs. The CAPs-IDD does not produce a total score; however, it was used in the present review to summarise the psychometric properties of the included measures.

There are two parts to the CAPs-IDD. Part 1 relates to the conceptual and measurement model of an instrument, describing basic information about the measure, how the measure was developed and measurement characteristics. Part 2 pertains to psychometric properties and summarises information about the validity (criterion, content, construct and face), reliability (internal consistency, reliability and measurement error), objectivity of application, objectivity of interpretation and feasibility of a measure. The review focussed on Part 2 of the CAPs-IDD framework (pertaining to psychometric properties) to discuss the findings of the quality appraisal. The results for Part 1 of the framework are reported in Tables S7–S15 in the supporting information. For the purpose of this review, measurement error was not discussed as no information pertaining to this was identified in the reviewed papers. Furthermore, face validity was not reviewed separately, as it overlapped with content validity in the CAPs-IDD framework, which encompassed both the relevance and comprehensiveness of items in the measures.

In order to determine the psychometric quality of included measures, they were subsequently rated on the four-point scale used by Flynn *et al.* (2017) (++) excellent; + good; – fair; - poor). Further information can be found in Table 2. With regard to interpreting results of factor analyses, a root mean square error of

approximation of $\leq .06$ and a comparative fit index of ≥ 0.95 were rated as a good fit (Hu & Bentler 1999). Samples of 100–200 participants are generally considered adequate for examining the reliability and validity of instruments (EFPA 2013), and therefore, samples with fewer than 100 participants were considered ‘inadequate’.

The second author (J. Y. L.) independently extracted data from the included papers and rated the quality of the measures. Differences in ratings were discussed until a consensus was reached.

Results

The screening process yielded 10 papers, which reviewed nine measures. The measures included in the review are listed in Table 3. Five papers presented initial studies of the psychometric properties of a measure (McGillivray *et al.* 2009; Wigham *et al.* 2011; Brooks *et al.* 2013; Chaplin *et al.* 2013; Raczka *et al.* 2020), and four papers detailed further analyses of measures, which had been previously validated in other studies (Devine *et al.* 2010; Rojahn *et al.* 2011; Briscoe *et al.* 2019; Wigham *et al.* 2021). Hall *et al.* (2014) reported both initial psychometric data of a measure and additional data from analyses of another measure. A description of each measure is presented in Table S5.

In summary, five measures included items that pertained to a broader spectrum of disorders or mental health difficulties (ADAMS, ADD, CORE-LD, Mini PAS-ADD and SAINT), two measured PTSD (IES-IDs and LANTS), and two measured QoL (Mini-MANS-LD and PWI-ID). Five of the measures reviewed were self-report (CORE-LD, IES-IDs, Mini-MANS-LD, PWI-ID and SAINT), and three were designed to be used with an informant (ADAMS, ADD and Mini-PAS-ADD). The LANTS included both a self-report and an informant-report scale. Three measures were recommended to detect changes over time and/or in response to an intervention (e.g., as a routine outcome measure; CORE-LD, Mini-MANS-LD and PWI-ID).

A summary of the psychometric data is presented in Table 4. The CAPs-IDD tables that provide a comprehensive description of the conceptual and measurement model and psychometric properties of the nine measures are presented in Tables S7–S15.

Table 2 Criteria used to interpret the results of psychometric evaluations

Measure	Range	Rating	Source
Correlation coefficient	.7–1	++	Hinkle <i>et al.</i> (2002)
	.5–.69	+	
	.3–.49	–	
	< .29	-	
Cronbach's alpha	0.9–1	++	Cicchetti (1994)
	0.8–0.89	+	
	0.7–0.79	–	
	0.0–0.69	-	
Intra-class correlations	.7–1	++	Cicchetti & Sparrow (1981)
	.5–.69	+	
	.3–.49	–	
	< .29	-	

Table 3 List of measures included in the present review.

Measure	Acronym	Original authors
Anxiety, Depression and Mood Scale	ADAMS	Esbensen <i>et al.</i> (2003)
Assessment of Dual Diagnosis	ADD	Matson & Bamburg (1998)
Clinical Outcomes in Routine Evaluation-Learning Disabilities	CORE-LD	Brooks <i>et al.</i> (2013)
Impact of Events Scale-Intellectual Disabilities	IES-IDs	Hall <i>et al.</i> (2014)
Lancaster and Northgate Trauma Scales	LANTS	Wigham <i>et al.</i> (2011)
Mini-Maslow Assessment of Needs Scale-Learning Disabilities	Mini-MANS-LD	Raczka <i>et al.</i> (2020); Skirrow & Perry (2009)
Mini Psychiatric Assessment Schedule for Adults with a Developmental Disability	Mini PAS-ADD	Moss (2002); Prosser <i>et al.</i> (1998)
Personal Wellbeing Index-Intellectual Disability	PWI-ID	Cummins <i>et al.</i> (2003); McGillivray <i>et al.</i> (2009)
Self-Assessment and Intervention (self-report section)	SAINT	Chaplin <i>et al.</i> (2013)

Sample characteristics

A summary of the sample characteristics is presented in Table S6. Most of the included studies were conducted in the UK ($n = 7$; 77.8%), in addition to one study that was conducted in Australia and another in the USA. Sample sizes ranged from 33 to 324. The percentage of male participants in each sample ranged from 40.7% to 85.9%. Seven studies (77.8%) reported the mean age of the participants, and this ranged between 33.0 and 45.6 years. The age range for participants was not reported in all of the papers, although they specified that only adults aged 18+ years were recruited. Participants were recruited from a range of sources, including clinical services for people with IDs, residential services and day centres. Participants' ethnicity was reported in four studies (44.4%), in which at least 71.1% of the samples were White. With regard to the severity of ID, one study included participants with severe and profound IDs (who made up <50% of the sample) and another included participants with borderline IDs (6.3% of the sample). The remaining studies only recruited participants with either a mild or moderate ID, although the breakdown of severity of ID within the sample was not always reported.

Reliability

Internal consistency

Internal consistency is the extent to which items in a questionnaire are correlated and therefore measure

the same concept (Terwee *et al.* 2007). Internal consistency was assessed for eight measures and was generally high. Three measures had excellent total score internal consistencies (ADAMS, ADD and IES-IDs), three measures were rated as having 'good' internal consistency (CORE-LD, LANTS and SAINT) and two were rated as 'fair' (Mini-MANS-LD and PWI-ID). However, subscale internal consistencies across the measures were generally lower, as detailed in Table 4.

Test-retest

Test-retest reliability refers to the degree to which repeated administrations of a measure provide similar responses (Terwee *et al.* 2007). The time period between administrations is often 1 or 2 weeks, to prevent recall while ensuring that clinical change has not occurred (Terwee *et al.* 2007). Test-retest reliability was reported for five measures (CORE-LD, IES-IDs, LANTS, SAINT and PWI-ID). The time period between administrations of the measure ranged from 1 to 6 weeks. The coefficients ranged between good (CORE-LD, LANTS informant scale and PWI-ID) and excellent (IES-IDs, LANTS self-report scale and SAINT).

Validity

Criterion validity

Criterion validity refers to the extent to which scores on an instrument relate to a 'gold standard' measure

Table 4 Data on the reliability and validity of measures included in the present review

Measure	Reliability			Validity			Sample size (N)
	Internal consistency	Test-retest	Criterion	Content	Construct		
ADAMS	Total score: $\alpha = 0.90$ (++) Mean subscale: $\alpha = 0.83$ (+) Subscales: $\alpha = 0.78-0.85$ (- to +)	ND	ND	ND	Confirmatory factor analysis (-): CFI = .89, RMSEA = .10 Convergent (++): ADAMS depression and ADD depressed mood subscales ($r_s = .77, p < .000$) ADAMS general anxiety and ADD anxiety subscales ($r_s = .75, p < .000$) ADAMS Manic/Hyperactive and ABC Hyperactivity subscales ($r_s = .75, p < .000$) Confirmatory factor analysis (-): Model would not converge after over 60 iterations Convergent (+ to ++): ADD depressed mood and ADAMS depression subscales ($r_s = .77, p < .000$) ADD anxiety and ADAMS general anxiety subscales ($r_s = .75, p < .000$) ADD Depression subscale and ABC Lethargy subscale ($r_s = .63, p < .000$) Convergent (+): CORE-OM ($r_s = .68, p < .001$)	263 (adequate)	
ADD	Total score: $\alpha = 0.91$ (++) Mean subscale: $\alpha = 0.59$ (-) Subscales: $\alpha = 0.18-0.84$ (- to +)	ND	ND	ND		263 (adequate)	
CORE-LD	$\alpha = 0.80$ (+)	1 week $n = 50$ $p = .64$ (+)	ND	Methodology supports content validity (items adapted from an established measure, focus group contributed to measure development)		Brooks et al. (2013): 324 (adequate) Briscoe et al. (2019): 41 (inadequate) 40 (inadequate)	
IES-IDs	Total score: $\alpha = 0.90-0.91$ (++) Subscales: $\alpha = 0.61-0.88$ (- to +)	2 weeks $n = 40$ Total score: ICC = .86 (++) Subscales: ICC = .65-.85 (+ to ++)	ND	Methodology supports content validity (items adapted from an established measure, mental health professionals were consulted during measure development)	Convergent (- to ++): GDS self-report scale ($r = .60, p < .001$) GDS informant scale ($p > .05$) GAS ($r = .40, p < .01$) LANTS self-report scale ($r = .76, p < .001$) LANTS informant subscales ($r = .28-.37, p < .05$) Trauma frequency ($r = .35, p < .05$)	Wigham et al. (2011): 99 (inadequate) Hall et al. (2014): 40 (inadequate)	
LANTS	Self-report scale: $\alpha = 0.84-0.89$ (+) Informant subscales: $\alpha = 0.80-0.92$ (+ to ++)	Wigham et al. (2011): 5-6 weeks $n = 48$ Self-report scale: $r_s = .72, p < .01$ (++) Informant subscales: $n = 33$ $r_s = .57-.59, p < .01$ (+) Hall et al. (2014): 2 weeks Self-report scale: $n = 40$ ICC = .92 (++) Informant subscales: $n = 32$ ICC = .80-.84 (++)	ND	Methodology supports content validity (items derived by consulting a number of sources e.g. literature reviews and focus groups)	Factor analysis of self-report scale (-): Four factor structure which accounted for 53.46% of the variance. Seven items removed due to particularly high skewness and/or kurtosis. Kaiser-Meyer-Olkin value was acceptable (0.75). Convergent: Self-report scale (- to ++): IES (general population measure) subscales ($r_s = .58-.62, p < .01$) IES-IDs ($r = .76, p < .001$) BLESID ($r_s = .45, p < .01$) BSI ($r_s = .62, p < .01$) Informant subscales (- to +): IES-IDs ($r = .28-.37, p < .05$) PEDS ($r_s = .35-.50, p < .01$) PAS-ADD ($r_s = .26-.62, p < .01$) BLESID ($r_s = .54-.64, p < .01$) Convergence between LANTS self-report and informant subscales (-): Behavioural changes ($r_s = .20, p < .005$) Frequency ($r_s = .13, p > .05$) Severity ($r_s = .16, p > .05$)	Wigham et al. (2011): 99 (inadequate) Hall et al. (2014): 40 (inadequate)	

Table 4. (Continued)

Measure	Reliability		Validity				Sample size (N)
	Internal consistency	Test-retest	Criterion	Content	Construct		
Mini-MANS-LD	$\alpha = 0.74$ (-)	ND	ND	Methodology supports content validity (items adapted from an established measure, experts by experience were consulted)	Convergent (+); PWI-ID ($r = .67, p < .001$)	33 (inadequate)	
Mini PAS-ADD	ND	ND	Sensitivity: 100% (informants and psychiatrist correctly identified the same five individuals who had a potential mental health problem) Specificity: 77% (informants and psychiatrist correctly identified 17 people who did not have mental health problem; psychiatrist did not identify a mental health problem for five people, whereas the informants using the Mini PAS-ADD scored them above the threshold)	Methodology supports content validity (items adapted from an established measure)	ND	96 (inadequate)	
PWI-ID	$\alpha = 0.76$ (-)	1-2 weeks $n = 31$ $r = .67, p < .05$ (+) ICC = .57 (+)	ND	ND	Convergent (- to -); Life as a whole: $r = .27-.44, p < .05$ Factor analysis: Two factor structure which accounted for 57.97% of the variance. However, as one of the items loaded onto both factors and the minimum number of items for a factor was not met, the analysis resolved to a single factor.	114 (adequate)	
SAINT	$\alpha = 0.83$ (+)	1 week $n = 20$ $r = 0.90, p < .01$ (++)	ND	Methodology supports content validity (Delphi methods and focus groups were employed to inform the contents of the measure)	Convergent (+ to +); GDS: $r = 0.71; p < .01$ GAS: $r = 0.58; p < .01$	54 (inadequate)	

Note: ++ Excellent; + Good; - Fair; - Poor.

Abbreviations: CFI, comparative fit index; ICC, intraclass correlation coefficient; ND, no data; RMSEA, root mean square error of approximation.

Measures: BLESID, Bangor Life Events Schedule for Intellectual Disabilities (Hulbert-Williams et al. 2011); BSI, Brief Symptom Inventory (Derogatis 1993); CORE-OM, Clinical Outcomes in Routine Evaluation – Outcome Measure (Evans et al. 2002); GAS, Glasgow Anxiety Scale; GDS, Glasgow Depression Scale; IES, Impact of Events Scale (Horowitz et al. 1979); PAS-ADD, Psychiatric Assessment Schedule for Adults with Developmental Disability; PEDS, Paediatric Emotional Distress Scale (Saylor et al. 1999).

(Terwee *et al.* 2007). Hatton & Taylor (2013) noted that an issue with testing validity stringently in this field is the lack of 'gold standard' measures of mental health problems for people with IDs. Clinical opinion is still preferred by many researchers as the gold standard method (Perez-Achiaga *et al.* 2009). In the present review, criterion validity was evaluated only for one measure (Mini PAS-ADD), by examining the sensitivity and specificity the measure compared with an assessment by a psychiatrist. Sensitivity analysis was found to be perfect; however, specificity analysis was lower.

Content validity

Content validity refers to the extent to which concepts are represented by the items in the measure (Terwee *et al.* 2007). Content validity is deemed to be good if a clear description of the concept being measured and item selection is provided, in addition to the target population and experts being involved in the measure development process (Terwee *et al.* 2007). In the present review, six measures addressed at least one aspect of content validity (CORE-LD, IES-IDs, LANTS, Mini-MANS-LD, Mini-PAS-ADD and SAINT). The items in four of these (CORE-LD, IES-IDs, Mini-MANS-LD and Mini-PAS-ADD) were derived from established measures of mental health or wellbeing. Four measures were reported to be developed through consultation with both people with IDs and mental health experts (CORE-LD, LANTS, Mini-MANS-LD and SAINT). Mental health professionals were consulted during the development of the IES-IDs.

Construct validity

Construct validity is the extent to which an instrument measures the distinct construct it is intending to measure (Markus & Lin 2010). Convergent validity, a subset of convergent validity, refers to the extent to which measures of theoretically related constructs converge, therefore suggesting that they capture a common construct (Carlson & Herdman 2012). Construct validity also refers to the structural validity of a measure and whether different dimensions within the measure correlate.

Convergent validity was examined in all measures except for the Mini-PAS-ADD. The CORE-LD and Mini-MANS-LD were correlated with only one other

measure, which may limit assessment of convergent validity, while the other measures were correlated with more than one measure. There was a broad range in the strength of significant correlations with other measures and five measures had a minimum rating of 'good' (ADAMS, ADD, CORE-LD, Mini-MANS-LD and SAINT). Six measures demonstrated excellent convergent validity with at least one other measure (ADAMS, ADD, IES-IDs, LANTS and SAINT). It appeared that the relationship between self-report and informant-report measures was poorer, as the correlation between the IES-IDs and the GDS informant scale was not significant, while the correlations between the IES-IDs and the LANTS informant subscales ranged from poor to fair. The correlation between the LANTS informant subscales and PAS-ADD ranged from poor to good. Furthermore, the convergence between the LANTS self-report and informant-report subscales was also poor, as the magnitude of the correlation with the behavioural changes subscale was low, while the correlations with the frequency and severity subscales were not significant.

There were insufficient investigations into the factorial structures of included measures. A factor analysis was attempted for four of the measures (ADAMS, ADD, LANTS self-report scale and PWI-ID). This was rated as 'poor' for the ADAMS, while the model would not converge for the ADD. The anticipated factor structures were confirmed for the PWI-ID and LANTS, although seven items were removed from the analysis of the LANTS due to particularly high skewness and/or kurtosis. There are different recommendations in the literature for the number of participants required for factor analytic techniques, for example, Guadagnoli & Velicer (1988) suggested $n = 100-200$, whereas a minimum of 1:5 item:case ratio was recommended by Floyd & Widaman (1995). With regard to the adequacy of sample sizes in the included studies that examined factor structure, this was considered acceptable for the ADAMS, ADD (both $n = 263$) and PWI-ID ($n = 114$). The sample size for the LANTS self-report study was slightly below the recommendation ($n = 98$).

Objectivity of application, interpretation, norming and fairness

With regard to the objectivity of application and interpretation, some instructions for the

administration of the CORE-LD and IES-IDs were reported in the published papers although instructions for coding were not. A short manual was available for administration and coding for the Mini-MANS-LD. A more comprehensive manual is available for the Mini-PAS-ADD and PWI-ID. Guidelines for administration and coding were not found through a web search for the remaining measures. With regard to the SAINT, the authors reported that it was not intended to be coded as it has not been designed as a diagnostic tool, and its purpose was to help individuals with IDs to recognise symptoms of mental distress.

In the included studies, little to no information was reported with regard to normative or comparative data from the general population. McGillivray *et al.* (2009) compared PWI-ID ratings with PWI ratings from general population samples and found that the total scores did not differ significantly. For included measures which were adapted from measures designed for the general population (for example, the CORE-LD and IES-IDs), normative data may be found in the published papers assessing the psychometric properties of the general population measures, for example the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM; Evans *et al.* 2002) and the Impact of Events Scale-Revised (IES-R; Weiss & Marmar 1997).

In terms of representativeness and generalisability, convenience samples were used in all of the included studies, which may have limited fairness concerning culture, gender and age. Data on ethnicity and age were not reported consistently in the studies. Further information may be found in the 'Sample characteristics' section of this review.

Feasibility

No information on the percentage of missing values was reported in the included papers. No information regarding the ease of administration, burden of completing the measure or acceptability was reported for the ADAMS or ADD. Briscoe *et al.* (2019) reported that participants indicated that they found completing the CORE-LD easier than the CORE-OM. In terms of the IES-IDs, professionals in an Adult Community Learning Disability Team were consulted to modify the language of the IES-R in order to ensure acceptability. An interviewer script

was developed to increase the ease of administration so that the IES-IDs could be administered as a semi-structured interview. The LANTS was developed via consultation with a clinical sample, carers, advocates and clinicians, to ensure acceptability and inclusiveness. The authors reported that administration took between 10 and 20 min to complete. The Mini-MANS-LD was also developed following consultation with a group of experts by experience. Accessibility was enhanced by using pictures and colour-coded faces as prompts. The authors reported that it was rated by participants as 'easy to use' and acceptable to people with IDs and that administration took on average less than 12 min to complete. Information on the acceptability was not reported for the Mini PAS-ADD; however, the authors reported that interviewers were provided with training on how to administer the measure and that finding appropriate time to administer the tool was identified as a difficulty by informants. The PWI-ID included a pre-testing protocol to enhance ease of administration, by identifying the level of complexity respondents were able to use the scale. Participants were given the choice of an 11-, 5-, 3- and 2-point scale. Administration took on average between 10 and 20 min. Finally, the developers of the SAINT consulted with professional experts and service user experts to enhance the acceptability and feasibility of the measure.

Discussion

The psychometric properties of nine measures were examined across the 10 papers considered. Although internal consistency was examined in eight of the nine measures, test-retest reliability was only assessed for five measures. Furthermore, criterion validity was only assessed for one measure. Six measures addressed at least one aspect of content validity. Convergent validity was examined for eight measures, although two measures were correlated with only one other measure. Factor analyses were attempted for four of the measures. Only three studies were considered to have an adequate sample size of more than 100 participants.

Based on the results of the present review, the CORE-LD, IES-IDs, LANTS and SAINT were deemed to have promising psychometric properties, as these measures had at least one rating of 'good'

across both dimensions of reliability and at least one dimension of validity. Additionally, these measures were developed through consultation with mental health professionals and/or people with IDs; thus, they were deemed to have good content validity. Although Hermans & Evenhuis (2010) previously suggested that the ADD was a promising informant-report measure, the present review indicates a lack of evidence on the quality of this measure for the mild to moderate adult ID population.

Several limitations and strengths of the studies that validated these four measures should be noted. With regard to the CORE-LD, Brooks *et al.* (2013) noted that the sample size meant that it was not possible to establish a cutoff score or to investigate whether the measure was more appropriate for some groups of people with ID and not others. Briscoe *et al.* (2019) reported that the strength of the correlation coefficient between the CORE-LD and CORE-OM was lower compared with the correlation between other related measures, for example the GAS-ID and the Becks Anxiety Inventory (Beck *et al.* 1988; Mindham & Espie 2003) or the GDS-LD and the Becks Depression Inventory (Beck *et al.* 1996; Cuthill *et al.* 2003). A strength of the CORE-LD was the emphasis on inclusivity and collaboration in the development of the measure. Brooks *et al.* (2013) reported receiving feedback from the individuals with ID who were involved developing the measure, such as 'I have enjoyed every minute of this research' and 'I felt valued' (p. 328). This is incredibly important given the barriers to participation in research that individuals with IDs face (Lennox *et al.* 2005).

In terms of the IES-IDs, Hall *et al.* (2014) reported that a limitation of their study was the small sample size, which meant that the factor structure of the measure could not be examined. However, the authors reported that a study strength was that the sample had experienced at least one traumatic event in their lives and were at risk of experiencing PTSD. This allowed an investigation into the whether there was a relationship between trauma frequency and symptomatology as measured by the IES-IDs, so that convergent validity could be assessed. The authors also contrasted the IES-IDs with the LANTS and noted that conceptually, the IES-IDs specifically assessed PTSD symptomology in response to a specific trauma, whereas the LANTS assessed more

general trauma-related psychopathology, in addition to symptoms of anxiety and depression, which are comorbid with PTSD.

Regarding the LANTS, Wigham *et al.* (2011, 2021) considered the inclusion of participants from both inpatient and community settings to be a study strength as this suggests that the LANTS may be utilised in both settings. Furthermore, the LANTS was developed via consultation with individuals with IDs, carers and clinicians, which supports their content validity. However, the sample recruited was 99% White British, which may limit the applicability for other ethnic groups. Wigham *et al.* (2011) also highlighted that the self-report version of the LANTS was only significantly correlated with one of the informant LANTS subscales and that the strength of this correlation was low. They posited that this may be because the two scales measured different aspects of trauma; the self-report version measured internal states whereas the informant scale measured observable behaviours. They suggested that construct validity was not compromised as both scales correlated significantly with the number of adverse life events experienced. This limitation is unlikely to be specific to this particular measure, and it has been suggested that the degree of convergence between self-report and informant-report scales may not reflect validity (Stancliffe 1995). This may be because informants cannot directly access the subjective experiences of individuals with IDs (Hartley & MacLean 2006). Although informant scales offer valuable information, it has been suggested that a mental health assessment of an individual with IDs should also include self-report questionnaires, which may add unique information about affective and cognitive symptoms which may not be apparent to caregivers (Mileviciute & Hartley 2015). A combination of these assessment approaches would enable clinicians to gain a more holistic understanding of an individual's presentation.

Chaplin *et al.* (2013) reported that a limitation of the SAINT was that test-retest reliability was assessed on a small proportion (37%) of the participants ($n = 20$) and that retest data were collected on the same day. A strength of the study was that the convergent validity was examined using the GDS-LD and GAS-ID, which were reported in previous systematic reviews to have promising psychometric properties. Furthermore, the SAINT was developed

through consultation with experts and service users. Chaplin *et al.* (2013) reported that an advantage of the SAINT was that it measured psychological distress more generally, rather than specific symptoms of depression and anxiety, which may present similarly in people with IDs. It is also unique, as the SAINT is not only a measure of distress but also covers specific coping strategies to reduce distress.

When considering the generalisability of the findings in the present review, it is important to note that the majority of participants were recruited from clinical or inpatient settings. For example, Briscoe *et al.* (2019) commented on their sample that comprised of forensic inpatients with IDs and other comorbidities. These samples may be unrepresentative of individuals with IDs in the general population as it is likely that these individuals experienced a higher level of distress compared with a community sample.

Limitations of the present review

Only studies that administered measures in the English language were included, and therefore, 33 articles that assessed the quality of measures administered in other languages were excluded. The decision was made to only include measures that were validated in the English language, to inform measurement choice for fellow English-speaking researchers and clinicians. When selecting measures for people with IDs in English-speaking countries, one may choose to translate measures that were validated in other languages, into English. However, it is recognised that cross-cultural adaptation of measures may be problematic for several reasons, such as the two languages having non-equivalent words, or items having very different meanings based on the specific cultural context (Epstein *et al.* 2015). Future researchers may wish to complete a review including reports on measures administered in languages other than English, to inform clinicians and researchers interested in selecting an appropriate measure in another language.

A further limitation was that the quality appraisal was somewhat limited by the lack of comprehensive evaluations of the psychometric properties of the included measures. There was very little information available on criterion and structural validity. Furthermore, little to no normative data from the

general population were reported and there was a lack of information on the time taken to complete the measures or how they were scored. It was therefore difficult to make comparisons between measures as not all aspects of reliability and validity were assessed for each measure. This highlights the need for continued research efforts into the quality of measures available for people with IDs.

Additionally, the agreement between reviewers on the quality of measures, which may have provided support for the validity of the quality assessments, was not calculated.

Implications for research and practice

Researchers and clinicians may use the findings of this review to make informed decisions when choosing a mental health or wellbeing measure for adults with mild to moderate IDs. Mileviciute & Hartley (2015) reported that self-report questionnaires may capture internalised experiences of people with IDs, which may not be apparent to carers. Therefore, although informant questionnaires offer valuable information, assessments should also include self-reported information. The CORE-LD, IES-IDs, LANTS and SAINT all include self-report scales. The CORE-LD may be used by clinicians and researchers interested in measuring the wellbeing, psychosocial functioning and emotional difficulties experienced by adults with mild to moderate IDs. The SAINT self-report section forms part of a guided self-help tool for people with IDs and may be used to assist people with IDs in recognising and reporting symptoms that indicate mental distress. Finally, the IES-IDs and LANTS may be used to screen for symptoms of PTSD. The LANTS also assesses for comorbid symptoms of anxiety and depression and, additionally, allows informants to provide information based on their observations of individuals with IDs.

Although two measures of QoL were identified in the present review, a lack of psychometrically sound measures of mental wellbeing, encompassing dimensions of hedonia and eudaimonia, for adults with mild to moderate IDs, was observed. The CORE-LD only included one positively worded item, 'Have you felt happy with the things you have done?' and so this tool may not be sufficient for those interested in measuring positive aspects of mental wellbeing. The results from this review therefore have

implications for research as they highlight a need to develop and validate measures of positive mental health, or mental wellbeing, in adults with mild to moderate IDs.

The present review also has implications for researchers who may be interested in adapting or evaluating measures for the ID population, as it has highlighted the impact of missing information on limiting psychometric research. Future researchers may wish to use the CAPs-IDD framework to guide their reporting and ensure that information pertaining to the development, administration, interpretation and feasibility of the measure is provided, in addition to the psychometric data. For example, many studies did not report how long a measure took to administer, which would be an important factor to consider when selecting a measure for this population. Additionally, providing a comprehensive description of the study sample, including a breakdown of the severity of ID, age range and ethnicity of participants, will enable readers to consider whether a measure may be suitable for a particular individual. Based on the findings of the present review, although internal consistency and convergent validity was often reported, fewer studies assessed test–retest reliability, criterion validity, content validity and structural validity. It is therefore recommended that future researchers endeavour to assess these dimensions of reliability and validity.

Conclusions

This review evaluated the psychometric properties of nine measures of common mental health problems and mental wellbeing in adults with mild to moderate IDs, administered in English. Four of these (CORE-LD, IES-IDs, LANTS and SAINT) were deemed to have promising psychometric properties. The results were limited by incomplete psychometric evaluations of measures, which made it difficult to compare measures. A paucity of psychometrically robust measures of mental wellbeing was observed. This review informs measurement choice for researchers and clinicians while highlighting a need for continued research efforts into the quality of measures available for people with IDs.

Acknowledgements

We thank UCL Library Services for their valuable support with the systematic literature search.

Source of funding

No funding was received for this study.

Conflict of interest

There are no conflicts of interests to declare.

Ethics approval

Not required for this study.

Authors' contributions

M. P. and K. S. designed the study. M. P. completed the literature search, extracted data and conducted the quality assessment. K. S. audited the screening process. J. Y. L. double rated the quality of the included measures. M. P. drafted the manuscript. All authors reviewed drafts of the manuscript.

Patient consent

Not required for this study.

Permission to reproduce material from other sources if any

No material was reproduced from other sources.

Data availability statement

The data that supports the findings of this study are available in the supporting information of this article.

References

- Aman M. G., Singh N. N., Stewart A. W. & Field C. J. (1985) The aberrant behavior checklist: a behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency* **89**, 485–91.
- American Psychiatric Association (2000) *Diagnostic and Statistical Manual of Mental Disorders, fourth edition text revision (DSM-IV-TR)*, Washington, DC.
- Beck A. T., Epstein N., Brown G. & Steer R. A. (1988) An inventory for measuring clinical anxiety: psychometric

- properties. *Journal of Consulting and Clinical Psychology* **56**, 893–7.
- Beck A. T., Steer R. A. & Brown G. K. (1996) Beck Depression Inventory. In: *The Psychological Corporation, 2nd Edition Manual* edn, San Antonio, TX.
- Beck A. T., Steer R. A. & Carbin M. G. (1988) Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clinical Psychology Review* **8**, 77–100.
- Braun V. & Clarke V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology* **3**, 77–101.
- Briscoe H., Ashworth S. & Shelton L. (2019) Establishing the concurrent validity of the Clinical Outcome in Routine Evaluation-Learning Disabilities scale. *Advances in Mental Health and Intellectual Disabilities* **13**, 237–44.
- Brooks M., Davies S. & Twigg E. (2013) A measure for feelings-using inclusive research to develop a tool for evaluating psychological therapy (Clinical Outcomes in Routine Evaluation-Learning Disability). *British Journal of Learning Disabilities* **41**, 320–9.
- Carlson K. D. & Herdman A. O. (2012) Understanding the impact of convergent validity on research results. *Organizational Research Methods* **15**, 17–32.
- Chaplin E., Chester R., Tsakanikos E., McCarthy J., Craig T. & Bouras N. (2013) Reliability and validity of the SAINT: a guided self-help tool for people with intellectual disabilities. *Journal of Mental Health Research in Intellectual Disabilities* **6**, 245–53.
- Cicchetti D. V. (1994) Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* **6**, 284–90.
- Cicchetti D. V. & Sparrow S. A. (1981) Developing criteria for establishing interrater reliability of specific items: applications to assessment of adaptive behavior. *American Journal of Mental Deficiency* **86**, 127–37.
- Cooke P. J., Melchert T. P. & Connor K. (2016) Measuring well-being: a review of instruments. *The Counseling Psychologist* **44**, 730–57.
- Cooper S.-A., Smiley E., Morrison J., Williamson A. & Allan L. (2007) Mental ill-health in adults with intellectual disabilities: prevalence and associated factor. *The British Journal of Psychiatry* **190**, 27–35.
- Creamer M., Bell R. & Failla S. (2003) Psychometric properties of the Impact of Event Scale-Revised. *Behaviour Research and Therapy* **41**, 1489–96.
- Cummins R. A. (1997) *Comprehensive Quality of Life Scale – Adult, Manual*, 5th edn, pp. 1–51. School of Psychology, Deakin University, Melbourne.
- Cummins R. A., Eckersley R., Pallant J., van Vugt J. & Misajon R. (2003) Development of a national index of subjective wellbeing: the Australian Unity Wellbeing Index. *Social Indicators Research* **64**, 159–90.
- Cuthill F. M., Espie C. A. & Cooper S. A. (2003) Development and psychometric properties of the Glasgow Depression Scale for people with a learning disability. Individual and carer supplement versions. *British Journal of Psychiatry* **182**, 347–53.
- Deci E. L. & Ryan R. M. (2008) Hedonia, eudaimonia, and wellbeing: an introduction. *Journal of Happiness Studies* **9**, 1–11.
- Derogatis L. R. (1993) *Brief Symptom Inventory: administration and scoring procedures manual*, 3rd edn. National Computer Systems, Minneapolis, MN.
- Devine M., Taggart L. & McLornian P. (2010) Screening for mental health problems in adults with learning disabilities using the Mini PAS-ADD Interview. *British Journal of Learning Disabilities* **38**, 252–8.
- Dunn K., Rydzewska E., Fleming M. & Cooper S. A. (2020) Prevalence of mental health conditions, sensory impairments and physical disability in people with co-occurring intellectual disabilities and autism compared with other people: a cross-sectional total population study in Scotland. *BMJ Open* **10**, e035280.
- Emerson E. & Hatton C. (2007) Mental health of children and adolescents with intellectual disabilities in Britain. *The British Journal of Psychiatry* **191**, 493–9.
- Epstein J., Santo R. M. & Guillemin F. (2015) A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. *Journal of Clinical Epidemiology* **68**, 435–41.
- Esbensen A. J., Rojahn J., Aman M. G. & Ruedrich S. (2003) Reliability and validity of an assessment instrument for anxiety, depression, and mood among individuals with mental retardation. *Journal of Autism and Developmental Disorders* **33**, 617–29.
- European Federation of Psychologists' Associations. (2013). *EFPA review model for the description and evaluation of psychological tests: test review form & notes for reviewers (version 4.2.6)*. Available at: <http://assessment.efpa.eu/documents/>
- Evans C., Connell J., Barkham M., Margison F., McGrath G., Mellor-Clark J. et al. (2002) Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *The British Journal of Psychiatry* **180**, 51–60.
- Floyd F. & Widaman K. F. (1995) Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment* **7**, 286–99.
- Flynn S., Vereenoghe L., Hastings R. P., Adams D., Cooper S. A., Gore N. et al. (2017) Measurement tools for mental health problems and mental well-being in people with severe or profound intellectual disabilities: a systematic review. *Clinical Psychology Review* **57**, 32–44.
- Guadagnoli E. & Velicer W. F. (1988) Relation of sample size to the stability of component patterns. *Psychological Bulletin* **103**, 265–75.
- Hagerty M. R., Cummins R., Ferriss A. L., Land K., Michalos A. C., Peterson M. et al. (2001) Quality of life indexes for national policy: review and agenda

- for research. *Bulletin of Sociological Methodology* **71**, 58–78.
- Hall J. C., Jobson L. & Langdon P. E. (2014) Measuring symptoms of post-traumatic stress disorder in people with intellectual disabilities: the development and psychometric properties of the Impact of Event Scale-Intellectual Disabilities (IES-s). *British Journal of Clinical Psychology* **53**, 315–32.
- Hartley S. L. & MacLean W. E., Jr. (2006) A review of the reliability and validity of Likert-type scales for people with intellectual disability. *Journal of Intellectual Disability Research* **50**, 813–27.
- Hatton C. & Taylor J. L. (2013) The assessment of mental health problems in adults with intellectual disabilities. In: *Psychological therapies for adults with intellectual disabilities* (eds J. L. Taylor, W. R. Lindsay, R. P. Hastings & C. Hatton), pp. 31–54. Wiley & Sons, Chichester.
- Hermans H. & Evenhuis H. M. (2010) Characteristics of instruments screening for depression in adults with intellectual disabilities: systematic review. *Research in Developmental Disabilities* **31**, 1109–20.
- Hermans H., van der Pas F. H. & Evenhuis H. M. (2011) Instruments assessing anxiety in adults with intellectual disabilities: a systematic review. *Research in Developmental Disabilities* **32**, 861–70.
- Hinkle D. E., Wiersma W. & Jurs S. G. (2002) *Applied statistics for the behavioural sciences*, 5th edn. Houghton Mifflin, Boston.
- Horowitz M., Wilner N. & Alvarez W. (1979) Impact of Event Scale: a measure of subjective stress. *Psychosomatic Medicine* **41**, 209–18.
- Hsieh K., Scott H. M. & Murthy S. (2020) Associated risk factors for depression and anxiety in adults with intellectual and developmental disabilities: five-year follow up. *American Journal on Intellectual and Developmental Disabilities* **125**, 49–63.
- Hu L. & Bentler P. (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* **6**, 1–55.
- Hulbert-Williams L., Hastings R. P., Crowe R. & Pemberton J. (2011) Self-reported life events, social support and psychological problems in adults with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities* **24**, 427–36.
- International Wellbeing Group (2006) *Personal Wellbeing Index*. Deakin University, Australian Centre on Quality of Life.
- Janssen R. & Maes B. (2013) Psychometric evaluation of a Dutch version of the Mini PAS-ADD for assessing psychiatric disorders in adults with different levels of intellectual disability. *Journal of Intellectual Disability Research* **57**, 689–702.
- Keyes C. L. M. (2002) The mental health continuum: from languishing to flourishing in life. *Journal of Health and Social Behavior* **43**, 207–22.
- Kovacs M. (1985) The Children's Depression Inventory (CDI). *Psychopharmacological Bulletin* **21**, 995–8.
- Lau A. L. D., Cummins R. A. & McPherson W. (2005) An investigation into the cross-cultural equivalence of the personal wellbeing index. *Social Indicators Research* **75**, 403–30.
- Lennox N., Taylor M., Rey-Conde T., Bain C., Purdie D. M. & Boyle F. (2005) Beating the barriers: recruitment of people with intellectual disability to participate in research. *Journal of Intellectual Disability Research* **49**, 296–305.
- Li C., Tsoi E. W., Zhang A. L., Chen S. & Wang C. J. (2013) Psychometric properties of self-reported quality of life measures for people with intellectual disabilities: a systematic review. *Journal of Developmental and Physical Disabilities* **25**, 253–70.
- Markus K. & Lin C. (2010) Construct validity. In: *Encyclopedia of research design* (ed. N. J. Salkind), pp. 230–3. SAGE Publications, Inc.
- Maslow A. (1943) A theory of human motivation. *Psychological Review* **50**, 370–96.
- Matson J. L. (1995) *The Matson evaluation of social skills for individuals with severe retardation*. Scientific Publishers, Inc., Baton Rouge, LA.
- Matson J. L. & Bamburg J. W. (1998) Reliability of the assessment of dual diagnosis (ADD). *Research in Developmental Disabilities* **19**, 89–95.
- Matson J. L., Gardner W. I., Coe D. A. & Sovner R. (1991) A scale for evaluating emotional disorders in severely and profoundly mentally retarded persons: Development of the Diagnostic Assessment for the Severely Handicapped (DASH) scale. *British Journal of Psychiatry* **159**, 404–9.
- Matson J. L., Kazdin A. E. & Senatore V. (1984) Psychometric properties of the psychopathology instrument for mentally retarded adults. *Applied Research in Mental Retardation* **5**, 81–9.
- McGillivray J. A., Lau A. L. D., Cummins R. A. & Davey G. (2009) The utility of the personal wellbeing index intellectual disability scale in an Australian sample. *Journal of Applied Research in Intellectual Disabilities* **22**, 276–86.
- Mileviciute I. & Hartley S. L. (2015) Self-reported versus informant-reported depressive symptoms in adults with mild intellectual disability. *Journal of Intellectual Disability Research* **59**, 158–69.
- Mindham J. & Espie C. A. (2003) Glasgow Anxiety Scale for people with an Intellectual Disability (GAS-ID): development and psychometric properties of a new measure for use with people with mild intellectual disability. *Journal of Intellectual Disability Research* **47**, 22–30.
- Moss S. (2002) *The Mini PAS-ADD interview pack*. Pavilion Pub.
- Moss S., Patel P., Prosser H., Goldberg D., Simpson N., Rowe S. et al. (1993) Psychiatric morbidity in older

- people with moderate and severe learning disability: I: development and reliability of the patient interview (PAS-ADD). *The British Journal of Psychiatry* **163**, 471–80.
- National Institute for Health and Care Excellence (NICE) (2011). *Common mental health disorders: identification and pathways to care*. Available at: <https://www.nice.org.uk/guidance/cg123>
- Novaco R. W. (1994) Anger as a risk factor for violence among the mentally disordered. In: *Violence mental disorder: developments in risk assessment* (eds J. Monahan & H. J. Steadman), pp. 21–59. University of Chicago Press, Chicago.
- Overall J. E. & Gorham D. R. (1962) The Brief Psychiatric Rating Scale. *Psychological Reports* **10**, 799–812.
- Perez-Achiaga N., Nelson S. & Hassiotis A. (2009) Instruments for the detection of depressive symptoms in people with intellectual disabilities: a systematic review. *Journal of Intellectual Disabilities* **13**, 55–76.
- Pouls K. P., Koks-Leensen M. C., Mastebroek M., Leusink G. & Assendelft W. (2021) Adults with intellectual disabilities and mental health disorders in primary care: a scoping review. *British Journal of General Practice* **72**, e168–78.
- Prosser H., Moss S., Costello H., Simpson N., Patel P. & Rowe S. (1998) Reliability and validity of the Mini PAS-ADD for assessing psychiatric disorders in adults with intellectual disability. *Journal of Intellectual Disability Research* **42**, 264–72.
- Raczka R., Theodore K. & Williams J. (2020) An initial validation of a new quality of life measure for adults with intellectual disability: the Mini-MANS-LD. *Journal of Intellectual Disabilities* **24**, 177–93.
- Raczka R., Williams J. & Theodore K. (2014) Pilot evaluation of routine outcome measures for use with adults with learning disabilities. *The Bulletin of the Faculty for People with Intellectual Disabilities* **12**, 60–5.
- Reid K. A., Smiley E. & Cooper S. A. (2011) Prevalence and associations of anxiety disorders in adults with intellectual disabilities. *Journal of Intellectual Disability Research* **55**, 172–81.
- Reiss S. (1988) *Reiss screen for maladaptive behavior: test manual*. IDS Publishing Corporation.
- Rojahn J., Matson J. L., Lott D., Esbensen A. J. & Smalls Y. (2001) The Behaviour Problems Inventory: An instrument for the assessment of self-injury, stereotyped behavior, and aggression/destruction in individuals with developmental disabilities. *Journal of Autism and Developmental Disorders* **31**, 577–88.
- Rojahn J., Rowe E. W., Kasdan S., Moore L. & van Ingen D. J. (2011) Psychometric properties of the aberrant behavior checklist, the anxiety, depression and mood scale, the assessment of dual diagnosis and the social performance survey schedule in adults with intellectual disabilities. *Research in Developmental Disabilities* **32**, 2309–20.
- Ross E. & Oliver C. (2002) The relationship between levels of mood, interest and pleasure and 'challenging behaviour' in adults with severe and profound intellectual disability. *Journal of Intellectual Disability Research* **46**, 191–7.
- Ross E. & Oliver C. (2003) Preliminary analysis of the psychometric properties of the mood, interest & pleasure questionnaire (MIPQ) for adults with severe and profound learning disabilities. *British Journal of Clinical Psychology* **42**, 81–93.
- Saylor C. F., Swenson C. C., Reynolds S. S. & Taylor M. (1999) The Pediatric Emotional Distress Scale: a brief screening measure for young children exposed to traumatic events. *Journal of Clinical Child Psychology* **28**, 70–81.
- Schalock R. L. & Keith K. D. (1993) *Quality of Life Questionnaire Manual*. IDS Publishing Corporation, Worthington, Ohio.
- Schmidt S., Muhlan H. & Power M. (2006) The EUROHIS-QOL 8-item index: psychometric results of a cross-cultural field study. *European Journal of Public Health* **16**, 420–8.
- Skevington S. M. & Böhnke J. R. (2018) How is subjective well-being related to quality of life? Do we need two concepts and both measures? *Social Science and Medicine* **206**, 22–30.
- Skirrow P. & Perry E. (2009) *The Maslow assessment of needs scales*. Mersey care NHS Trust, Liverpool.
- Smith J. A. (2003) *Qualitative Psychology: A Practical Guide to Research Methods*. Sage Publications, Inc.
- Stancliffe R. J. (1995) Assessing opportunities for choice-making: a comparison of self- and staff reports. *American Journal on Mental Retardation* **99**, 418–29.
- Stancliffe R. J. & Parmenter T. R. (1999) The choice questionnaire: a scale to assess choices exercised by adults with intellectual disability. *Journal of Intellectual and Developmental Disability* **24**, 107–32.
- Stewart-Brown S., Samaraweera P. C., Taggart F., Kandala N. B. & Stranges S. (2015) Socioeconomic gradients and mental health: implications for public health. *The British Journal of Psychiatry* **206**, 461–5.
- Terwee C. B., Bot S. D., de Boer M. R., van der Windt D. A., Knol D. L., Dekker J. *et al.* (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology* **60**, 34–42.
- Townsend-White C., Pham A. N. T. & Vassos M. V. (2012) A systematic review of quality of life measures for people with intellectual disabilities and challenging behaviours. *Journal of Intellectual Disability Research* **56**, 270–84.
- Trompeter H. R., de Kleine E. & Bohlmeijer E. T. (2017) Why does positive mental health buffer against psychopathology? An exploratory study on self-compassion as a resilience mechanism and adaptive

- emotion regulation strategy. *Cognitive Therapy and Research* **41**, 459–68.
- Weiss D. S. & Marmar C. R. (1997) The Impact of Event Scale-Revised. In: *Assessing Psychological Trauma and PTSD: A Handbook for Practitioners* (eds J. P. Wilson & T. M. Keane), pp. 399–411. Guilford Press, New York, NY.
- WHOQOL Group (1995) The World Health Organization Quality of Life Assessment (WHOQOL): position paper from the World Health Organization. *Social Science and Medicine* **41**, 1403–9.
- Wigham S., Hatton C. & Taylor J. L. (2011) The Lancaster and Northgate Trauma Scales (LANTS): the development and psychometric properties of a measure of trauma for people with mild to moderate intellectual disabilities. *Research in Developmental Disabilities* **32**, 2651–9.
- Wigham S., Hatton C. & Taylor J. L. (2021) Short report and initial evaluation of the factor structure of the Lancaster and Northgate Trauma Scales (LANTS). *Research in Developmental Disabilities* **112**, 103914.
- Wille N., Badia X., Bonsel G., Burström K., Cavrini G., Devlin N. *et al.* (2010) Development of the EQ-5D-Y: a child friendly version of the EQ-5D. *Quality of Life Research* **19**, 875–86.
- World Health Organization (WHO) (2018) *International Classification of Diseases for Mortality and Morbidity Statistics*, 11th Revision edn.
- Zeilinger E. L., Nader I. W., Brehmer-Rinderer B., Koller I. & Weber G. (2013) CAPs-IDD: characteristics of assessment instruments for psychiatric disorders in persons with intellectual developmental disorders. *Journal of Intellectual Disability Research* **57**, 737–46.

Accepted 20 January 2023

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article.

Data S1. Supporting Information